On solving mixed integer non linear programs with separable non convexities

Claudia D'Ambrosio
(joint works with A. Frangioni, C. Gentile, J. Lee, A. Wächter)

CNRS & École Polytechnique, France

Discrete Optimization and Machine Learning 2019
30 July 2019

Outline

Outline

The class of MINLP problems

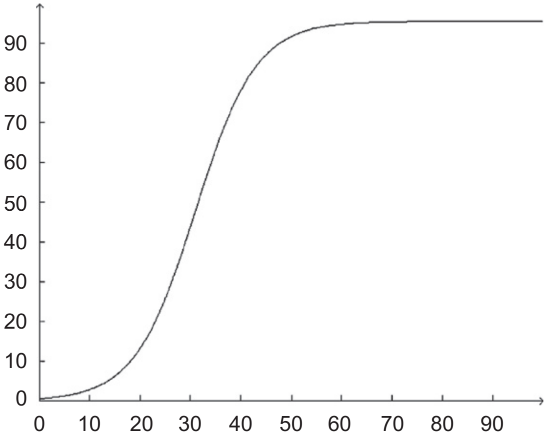
$$
\begin{aligned}
\min \sum_{j \in N} & C_j x_j \\
& f_i(x) + \sum_{k \in H_i} g_{ik}(x_k) \le 0 && \forall i \in M \\
& L_j \le x_j \le U_j && \forall j \in N \\
& x_j \text{ integer} && \forall j \in I
\end{aligned}
$$

where:

- $f_i : \mathbb{R}^n \to \mathbb{R}$ are convex functions $\forall i \in M$,
- $g_{ik} : \mathbb{R} \to \mathbb{R}$ are non convex univariate function $\forall i \in M, \forall k \in H_i$,
- $H_i \subseteq N \;\; \forall i \in M$,
- $I \subseteq N$, and
- L_j and U_j are finite $\forall i \in M, j \in H_i$

Example of application in Machine Learning

Model trained neural networks with sigmoid activation function

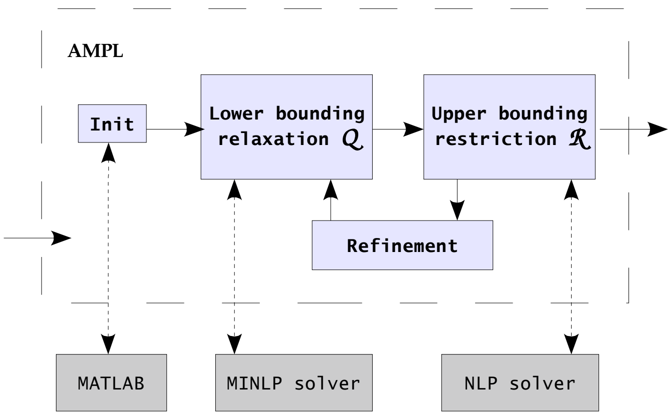


Outline

Outline

General Framework

Global optimization algorithm proposed in D'A., Lee, and Wächter (2009, 2012).



Outline

The Upper Bounding problem

Upper Bound of the original problem:

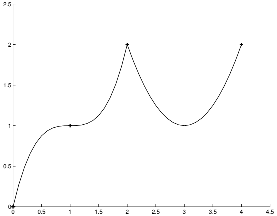
1. The integer variables are fixed;
2. We solve the resulting non convex NLP problem to local optimality;

$$
\begin{aligned}
\min \sum_{j \in N} & C_j x_j \\
& f_i(x) + \sum_{k \in H_i} g_{ik}(x_k) \leq 0 && \forall i \in M \\
& L_j \leq x_j \leq U_j && \forall j \in N \\
& x_j = \underline{x}_j && \forall j \in I
\end{aligned}
$$

Outline

The Lower Bounding problem: step 1

For simplicity, let us consider a term of the form $g(x_k) := g_{ik}(x_k)$: $g : \mathbb{R} \to \mathbb{R}$ is a univariate non convex function of x_k , for some k $(1 \leq k \leq n)$.



Automatically detect the concavity/convexity intervals or piecewise definition:

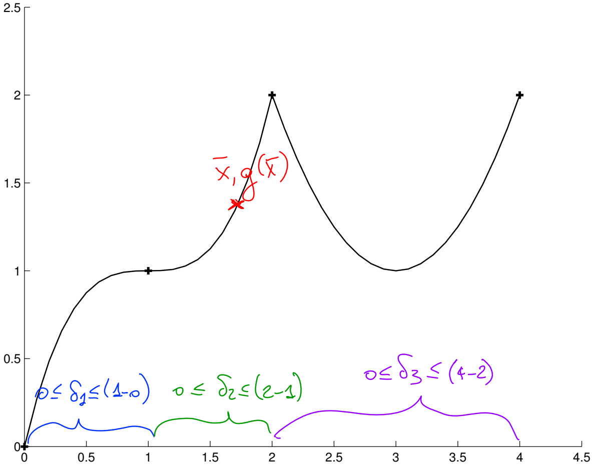
$[P_{p-1}, P_p] :=$ the p-th subinterval of the domain of g ($p \in \{1 \ldots \overline{p}\}$);

$\check{H} :=$ the set of indices of subintervals on which g is convex;

$\hat{H} :=$ the set of indices of subintervals on which g is concave.

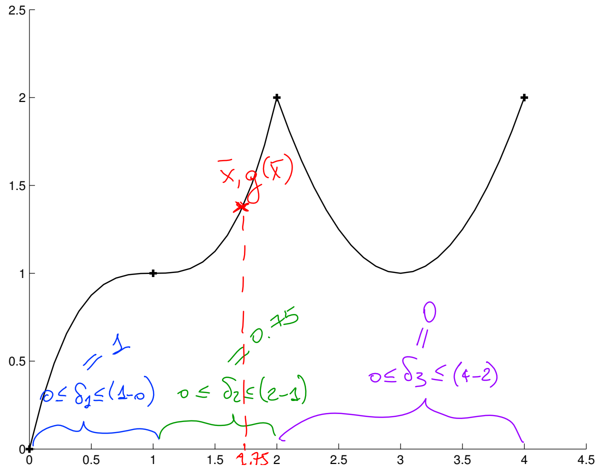
The Lower Bounding problem: step 2

Introduction of additional variables $\delta_p \in [0, P_p - P_{p-1}]$ such that $x_k = P_0 + \sum_{p=1}^{\bar{p}} \delta_p$



The Lower Bounding problem: step 2

Introduction of additional variables $\delta_p \in [0, P_p - P_{p-1}]$ such that $x_k = P_0 + \sum_{p=1}^{\bar{p}} \delta_p = 0 + 1 + 0.75 + 0$



The Lower Bounding problem: step 2

- All the δ's but at most 1 take either the lower or the upper bound value
- To model such behavior additional binary variables are needed: $z_p \in \{0,1\}\ \forall p$
- $z_1 \geq z_2 \geq \cdots \geq z_p$
- $\delta_p = \begin{cases} 0 & z_{p-1} = 0 \\ [0, P_p - P_{p-1}] & z_{p-1} = 1 \text{ and } z_p = 0 \\ P_p - P_{p-1} & z_p = 1 \end{cases}$

δ	$P_1 - P_0$	$P_2 - P_1$	$\dots$	$P_{p-1} - P_{p-2}$	$[0, P_p - P_{p-1}]$	0	$\dots$	0
z	1	1	$\dots$	1	0	0	$\dots$	0

The Lower Bounding problem: step 2

Replace the term $g(x_k)$ with:

$$\sum_{p=1}^{\bar{p}} g(P_{p-1} + \delta_p) - \sum_{p=1}^{\bar{p}-1} g(P_p) \ ,$$

and we include the following set of new constraints:

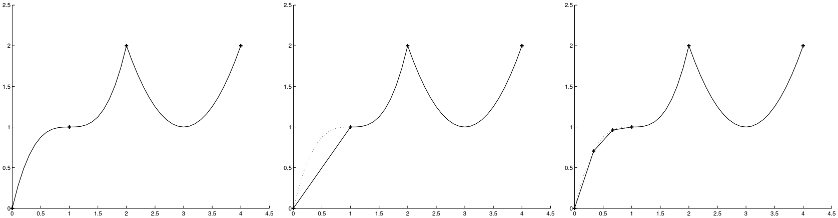
$$x_k = P_0 + \sum_{p=1}^{\bar{p}} \delta_p \ ;$$
$$\delta_p \geq (P_p - P_{p-1}) z_p \ , \ \forall p \in \check{H} \cup \hat{H} \ ;$$
$$\delta_p \leq (P_p - P_{p-1}) z_{p-1} \ , \ \forall p \in \check{H} \cup \hat{H} \ ;$$
$$0 \leq \delta_p \leq P_p - P_{p-1}, \ \forall p \in \{1, \ldots, \bar{p}\};$$

with two dummy variables $z_0 := 1$ and $z_{\bar{p}} := 0$ and two new sets of variables z_p (binary) and δ_p (continuous).

The Lower Bounding problem: step 3

Still non convex;

Use piece-wise linear approximation for the concave intervals:



The Lower Bounding problem: the convex MINLP model

Replace the term $g(x_k)$ with:

$$\sum_{p\in\check{H}} g(P_{p-1}+\delta_p)+\sum_{p\in\hat{H}}\sum_{b\in B_p} g(X_{p,b})\,\alpha_{p,b}-\sum_{p=1}^{\overline{p}-1} g(P_p)\ ,$$

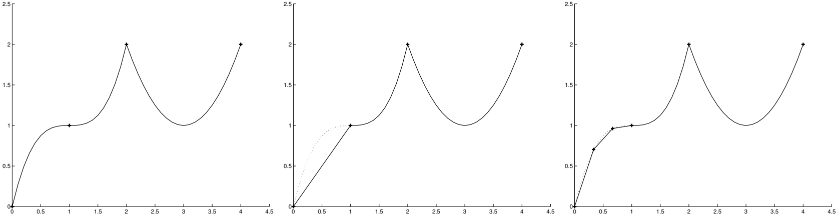
and we include the following set of new constraints:

$$\begin{aligned}
&P_0+\sum_{p=1}^{\overline{p}}\delta_p-x_k=0\ ;\\
&\delta_p-(P_p-P_{p-1})z_p\geq 0\ ,\ \forall p\in\check{H}\cup\hat{H}\ ;\\
&\delta_p-(P_p-P_{p-1})z_{p-1}\leq 0\ ,\ \forall p\in\check{H}\cup\hat{H}\ ;\\
&P_{p-1}+\delta_p-\sum_{b\in B_p} X_{p,b}\,\alpha_{p,b}=0\ ,\ \forall p\in\hat{H}\ ;\\
&\sum_{b\in B_p}\alpha_{p,b}=1\ ,\ \forall p\in\hat{H}\ ;\\
&\{\alpha_{p,b}:b\in B_p\}:=\text{SOS2}\ ,\qquad \forall p\in\hat{H}\ .
\end{aligned}$$

with two dummy variables $z_0:=1$, $z_{\overline{p}}:=0$ and the new set of variables $\alpha_{p,b}$.

Outline

Refining the Lower Bounding problem



- Add a breakpoint where the solution of problem $\mathcal{Q}$ of the previous iteration lies (global convergence);
- Add a breakpoint where the solution of problem $\mathcal{R}$ of the previous iteration lies (speed up the convergence).

Outline

Limits

- Solving the Lower Bounding problem can be **time consuming**

Limits

- Solving the Lower Bounding problem can be **time consuming**
- At each iteration we solve the Lower Bounding problem **from scratch**

Limits

- Solving the Lower Bounding problem can be **time consuming**
- At each iteration we solve the Lower Bounding problem **from scratch**
- **Large number of iterations** needed to converge

Lower Bounding problem tightening

Let us consider the convex pieces:

$$g(P_{p-1} + \delta_p) - g(P_{p-1})$$

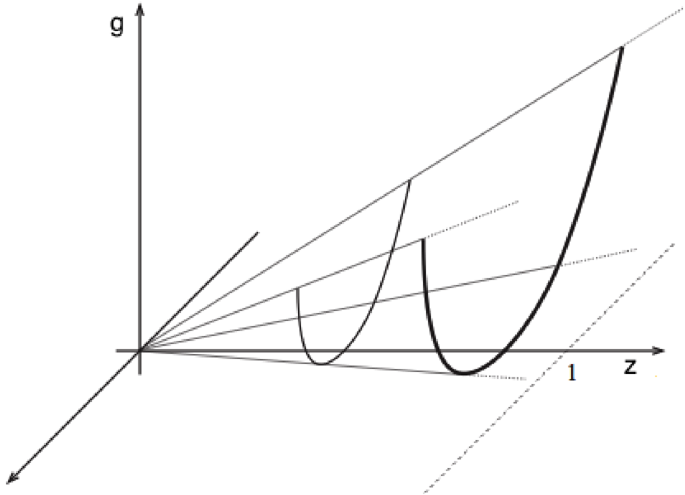
with

- $0 \leq \delta_p \leq (P_p - P_{p-1})z_{p-1}$
- $z_{p-1} \in \{0, 1\}$

Its **convex envelope** is:

$$z_{p-1}(g(P_{p-1} + \delta_p / z_{p-1}) - g(P_{p-1}))$$

Perspective function



Where can we exploit it?

Use it to solve the Lower Bounding problem:

- Reformulate the convex MINLP
- Stronger the convex continuous relaxation
- Generate stronger linear cuts
- Solve the convex MINLP with cutting plane

Outline

Computational Results

- `PC`: linearization of PR of LB problem
- `STD`: linearization of original LB problem
- `Bonmin`
- `Minotaur`
- `SCIP`

Tests on non linear knapsack problem and uncapacitated facility location problem.

10,000 seconds time limit.

The non linear knapsack problem

$$
\text{(NCK)} \quad \begin{array}{lll}
\max & \sum_{j \in N} p_j & \\
 & p_j - \frac{c_j}{1 + b_j \exp(-a_j(x_j + d_j))} \leq 0 & j \in N \\
 & \sum_{j \in N} x_j \leq C & \\
 & 0 \leq x_j \leq U & j \in N
\end{array}
$$

- $|N| \in \{10, 20, 50, 100, 200, 500\}$ (10 instances for each)
- Random Uniformly:
  $a_j \in [0.1\,, 0.2]$, $b_j \in [0\,, 100]$, $c_j \in [0\,, 100]$, and $d_j \in [-100\,, 0]$.

Results on the non linear knapsack problem

	Bonmin							Minotaur					
size	B-BB		B-OA		B-Hyb		B-OA-C	BNB-I		QG-I		QPD-I	
	time	gap	time	gap	time	gap	time	time	gap	time	gap	time	gap
10	1.06	-	0.25	-	0.59	-	0.27	0.22	-	0.11	-	0.09	-
20	2.99	-	0.34	-	2.12	-	0.32	0.53	-	0.22	-	0.16	-
50	13.8	-	0.65	-	8.05	-	0.62	2.97	-	1.07	-	0.63	-
100	78.9	-	9.16	-	7936	1.00	1.07	13.0	-	4.25	-	3.44	-
200	1000	-	5035	0.62	4019	0.88	2.24	88.5	-	37.8	-	28.6	-
500	tl	0.12	8035	0.62	9027	1.49	8.41	8621	0.07	7080	0.15	7692	0.16

Table: NCK: Bonmin and Minotaur options comparison

Results on the non linear knapsack problem

size	PC		STD		Bonmin	MINOTAUR			SCIP
	time	cuts	time	cuts	time	time	gap	bgap	time
10	0.014	96	0.015	102	0.267	0.09	-	-	0.07
20	0.021	155	0.019	195	0.324	0.16	-	-	0.10
50	0.048	431	0.085	678	0.617	0.63	-	-	0.21
100	0.072	947	0.183	1182	1.067	3.44	-	-	0.66
200	0.105	1780	0.565	2461	2.237	28.6	-	-	131.2
500	0.380	4681	3.593	7821	8.406	7080	0.15	0.05	181.4

Table: NCK: comparison among the different algorithms

The uncapacitated facility location problem

$$
\text{(UFL)} \quad
\begin{array}{lll}
\min & \sum_{k\in K} C_k y_k + \sum_{t\in T}\sum_{k\in K} s_{kt} & \\
 & g_{kt}(w_{kt}) - s_{kt} \leq 0 & t\in T \;,\; k\in K \\
 & \sum_{k\in K} w_{kt} = 1 & t\in T \\
 & 0 \leq w_{kt} \leq y_k & t\in T \;,\; k\in K \\
 & y_k \in \{0,1\} & k\in K
\end{array}
$$

For each combination $(|K|,|T|) \in \{(6,12),(12,24),(24,48)\}$ we generated 3 instances of increasing difficulty.

Results on the uncapacitated facility location problem

	Bonmin						Minotaur					
instance	B-BB		B-OA-C		B-Hy-C		BNB		QPD		QG-I	
	time	gap	time	gap	time	gap	time	gap	time	gap	time	gap
6x12x1	176	-	1.76	-	1.37	-	538	-	24.8	-	4.66	-
6x12x2	tl	1.16	7.25	-	5.64	-	tl	29.17	tl	51.08	65.5	-
6x12x3	tl	657.6	tl	∞	tl	∞	tl	∞	tl	315.5	tl	260.3
12x24x1	1592	-	9.68	-	7.14	-	tl	8.07	tl	66.57	57.4	-
12x24x2	tl	18.77	93.8	-	57.9	-	tl	∞	tl	∞	tl	17.40
12x24x3	tl	∞	tl	∞	tl	∞	tl	∞	tl	∞	tl	271.6
24x48x1	tl	84.70	116	-	132	-	tl	∞	tl	∞	2844	-
24x48x2	tl	73.44	tl	∞	tl	∞	tl	∞	tl	∞	tl	31.49

Table: UFL: Comparison among Bonmin and Minotaur options

Results on the uncapacitated facility location problem

instance	PC				STD				Bonmin			Minotaur		
	time	gap	bgap	cuts	time	gap	bgap	cuts	time	gap	bgap	time	gap	bgap
6x12x1	0.35	-	-	1673	0.26	-	-	1531	1.37	-	-	4.66	-	-
6x12x2	0.45	-	-	1842	0.42	-	-	1796	5.64	-	-	65.6	-	-
6x12x3	7921	-	-	33417	tl	54.3	52.4	180561	tl	657	796	tl	260	615
12x24x1	3.36	-	-	9565	2.55	-	-	8971	7.14	-	-	57.4	-	-
12x24x2	46.1	-	-	19653	27.3	-	-	17384	57.9	-	-	tl	17.4	10.5
12x24x3	tl	23.9	23.9	127380	tl	121	134	284557	tl	∞	1524	tl	272	1447
24x48x1	261	-	-	81372	316	-	-	102160	116	-	-	2844	-	-
24x48x2	tl	5.93	5.67	164809	tl	9.66	9.66	409177	tl	73.4	26.4	tl	31.5	24.6

Table: UFL: Comparison among different algorithms

Outline

Future Directions

With J. Lee, D. Skipper, D. Thomopulos

- Use **disjunctive cuts** to tighten the formulation instead of adding breakpoints

Future Directions

With J. Lee, D. Skipper, D. Thomopulos

- Use **disjunctive cuts** to tighten the formulation instead of adding breakpoints

With C. Artigues, A. Frangioni, C. Gentile, R. Trindade, S. Ulrich Ngueveu

- Is the LB problem formulation the **tightest**?
- Use the “**Piecewise linear bounding** of univariate functions” for the concave part

Future Directions

With J. Lee, D. Skipper, D. Thomopulos

- Use **disjunctive cuts** to tighten the formulation instead of adding breakpoints

With C. Artigues, A. Frangioni, C. Gentile, R. Trindade, S. Ulrich Ngueveu

- Is the LB problem formulation the **tightest**?
- Use the "**Piecewise linear bounding** of univariate functions" for the concave part

Thanks!